APPROXIMATION NUMERIQUE ET OPTIMISATION

G. ALLAIRE

16 Octobre 2018

CHAPITRE IV (début)

- ☞ Définitions, notations.
- ☞ Analyse convexe.
- ☞ Différentiabilité.
- ☞ Conditions d'optimalité: cas convexe.
- ☞ Conditions d'optimalité: cas non convexe.

DEFINITIONS

Soit V un espace vectoriel normé (on suppose que $dim(V) < \infty$ mais certains résultats restent vrais pour des espaces de Hilbert ou de Banach avec $dim(V) = \infty$.)

Soit $K \subset V$ un sous-ensemble non-vide. Soit $J : V \to \mathbb{R}$. On considère

$$\inf_{v \in K \subset V} J(v).$$

Définition. On dit que u est un minimum local de J sur K si

$$u \in K \quad \text{et} \quad \exists \delta > 0 \;, \; \forall v \in K \;, \; \|v - u\| < \delta \Longrightarrow J(v) \geq J(u) \;.$$

On dit que u est un minimum global de J sur K si

$$u \in K \quad \text{et} \quad J(v) \geq J(u) \quad \forall v \in K \;.$$

(différence: théorie $\leftrightarrow$ global / numérique $\leftrightarrow$ local)

Définition. Une suite minimisante du critère J sur l'ensemble K est une suite $(u^n)_{n\in\mathbb{N}}$ telle que

$$\forall\, n,\ u^n \in K \quad \text{et} \quad \lim_{n\to+\infty} J(u^n) = \inf_{v\in K} J(v).$$

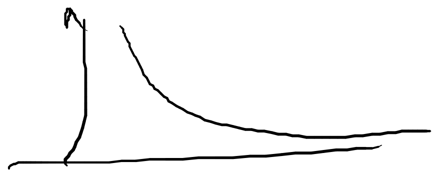
Par définition de l'infimum de J sur K **il en existe toujours !**

Remarque. La notation $\inf_{v\in K} J(v)$ désigne aussi un nombre, la valeur infimum, définie par

$$\left(\inf_{v\in K} J(v)\right) = \sup\{C \in \mathbb{R} \text{ tel que } C \le J(v)\ \forall\, v \in K\}$$

Par convention, $\inf_{v\in K} J(v) = -\infty$ si J n'est pas minorée sur K.

exemple $J(x) = \frac{1}{x}$ sur $\mathbb{R}^+_*$



$\inf_{\mathbb{R}^+_*} J(x) = 0$

Notation (pas universelle).

Si on écrit

$$\min_{v \in K} J(v)$$

c'est que l'on sous-entend que le minimum est atteint

$$\exists v^* \in K \text{ tel que } J(v^*) = \min_{v \in K} J(v)$$

Tandis que si l'on écrit

$$\inf_{v \in K} J(v)$$

on ne sait pas si le minimum est atteint ou pas.

Optimisation en dimension finie $V = \mathbb{R}^N$

Théorème. Soit K un ensemble fermé non vide de $\mathbb{R}^N$, et J une fonction continue sur K à valeurs dans $\mathbb{R}$ vérifiant la propriété, dite "infinie à l'infini",

$$\forall (u^n)_{n\geq 0} \text{ suite dans } K\ ,\ \lim_{n\to+\infty} \|u^n\| = +\infty \Longrightarrow \lim_{n\to+\infty} J(u^n) = +\infty\ .$$

Alors il existe au moins un point de minimum de J sur K. De plus, on peut extraire de toute suite minimisante de J sur K une sous-suite convergeant vers un point de minimum sur K.

(Idée: les fermés bornés sont compacts en dimension finie.)

Remarque. Si K est borné, alors la condition "infinie à l'infini" est vide.

Démonstration

Soit une suite minimisante $u_n \in K$

Il existe $c > 0$ tel que $\|u_n\| \leq c$ $\forall n$. En effet, sinon $\exists$ sous-suite $u_{n'}$ tq $\|u_{n'}\| \to +\infty$. A cause de l'hyp. "$\infty \in l'\infty$" $\Rightarrow J(u_{n'}) \to +\infty$ contradiction avec le caractère minimisant de la suite u_n

$\Rightarrow$ $\|u_n\| \leq c$ $\Rightarrow$ $\inf\limits_{v \in K} J(v) \equiv \inf\limits_{v \in K \cap B(0,c)} J(v)$

boule fermé de rayon c

$K \cap B(0,c)$ est un fermé borné et J continue

$\Downarrow$

J atteint son minimum sur $K \cap B(0,c)$ donc sur K

Optimisation en dimension infinie

Problème: une fonction continue sur un fermé borné n'atteint pas toujours son minimum !

Exemple de non-existence: Soit l'espace de Hilbert (de dimension infinie) des suites de carré sommable dans $\mathbb{R}$

$$\ell_2(\mathbb{R}) = \left\{ x = (x_i)_{i\geq 1} \text{ tel que } \sum_{i=1}^{+\infty} x_i^2 < +\infty \right\},$$

muni du produit scalaire $\langle x, y\rangle = \sum_{i=1}^{+\infty} x_i y_i$ et de la norme associée $\|x\| = \sqrt{\langle x, x\rangle}$. On considère le problème

$$\inf_{x\in\ell_2(\mathbb{R})} \left\{ J(x) = \left(\|x\|^2 - 1\right)^2 + \sum_{i=1}^{+\infty} \frac{x_i^2}{i} \right\}$$

qui n'admet pas de point de minimum.

$J(x) = \left(\|x\|^2 - 1\right)^2 + \sum_{i \geq 1} \frac{x_i^2}{i} \geqslant 0$

Démonstration

On a $\inf_{\ell^2(\mathbb{R})} J(x) \equiv 0$! En effet, soit la suite (minimisante) $x^n = (0 - 0, 1, 0 \dots 0 \dots)$ (position n)

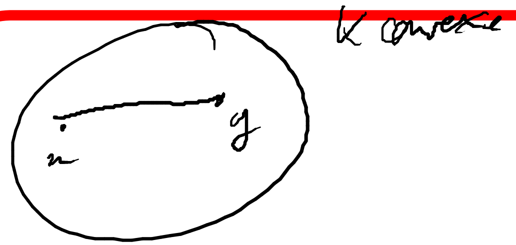
$\Rightarrow \|x^n\| = 1$

$J(x^n) = \frac{1}{n} \longrightarrow 0 \quad n \to +\infty$

$\exists$? x^* tq $J(x^*) = 0$ si oui, alors $\|x^*\| = 1$ et $x_i^* = 0 \quad \forall i$

Contradiction

Donc l'inf n'est pas atteint !

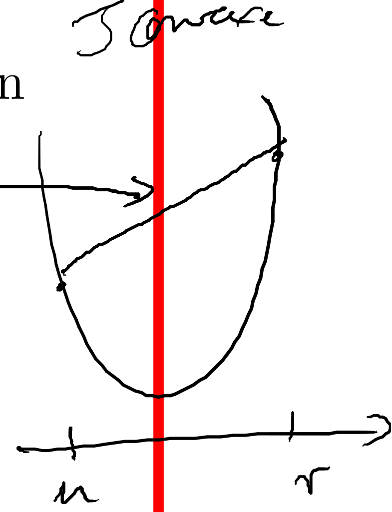


Analyse convexe

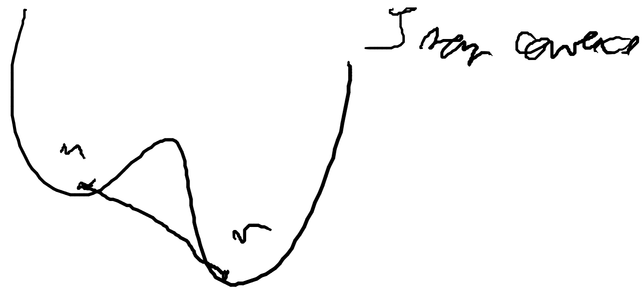
Définition. Un ensemble $K \subset V$ est dit **convexe** si, pour tout $x, y \in K$ et tout réel $\theta \in [0,1]$, l'élément $(\theta x + (1-\theta)y)$ appartient à K.

Définition. On dit qu'une fonction J définie sur un ensemble convexe non vide $K \in V$ et à valeurs dans $\mathbb{R}$ est **convexe** sur K si et seulement si

$$J(\theta u + (1-\theta)v) \leq \theta J(u) + (1-\theta)J(v) \quad \forall\, u, v \in K \;,\; \forall\, \theta \in [0,1] \;.$$



De plus, J est dite **strictement convexe** si l'inégalité est stricte lorsque $u \neq v$ et $\theta \in]0,1[$.



Unicité

Proposition. Si J est strictement convexe, alors il existe au plus un point de minimum.

Preuve. en effet supposons que $\exists\, u, v \in K$ tq $J(u) = J(v) = \min_K J$

$$w = \frac{1}{2}(u+v) \qquad J(w) < \underbrace{\frac{1}{2} J(u) + \frac{1}{2} J(v)}_{= \min_K J} \quad \text{si } u \neq v$$

contradiction

$$\Rightarrow \boxed{u = v}$$

Convexe $\Rightarrow$ local=global

Proposition. Si J est une fonction convexe sur un ensemble convexe K, tout point de minimum local de J sur K est un minimum global.

Preuve. Soit $u \in K$ point de min local

Soit $v \in K$ ———— global

$w = (1-\theta)u + \theta v = u + \theta(v-u)$ $\quad \theta \geq 0$ "petit"

$J(w) \leq (1-\theta)J(u) + \theta J(v)$

si θ petit $\Rightarrow$ $\|w-u\| = \theta\|v-u\| \leq \delta$ dans la déf. du min local

$\Rightarrow$ $J(u) \leq J(w)$

$\leq (1-\theta)J(u) + \theta J(v)$

$\Rightarrow$ $\theta J(u) \leq \theta J(v)$ $\Rightarrow$ $J(u) \leq J(v) = \min_K J$

$\Rightarrow$ $\boxed{J(u) = \min_K J}$

Conditions d'optimalité

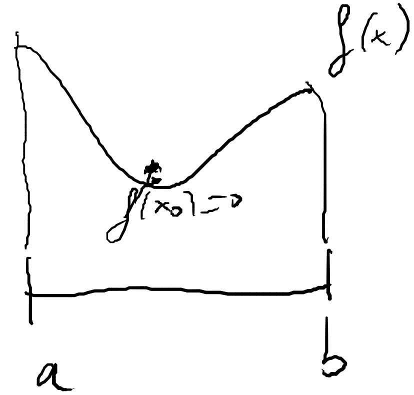
Pour caractériser ou calculer les solutions de

$$\inf_{v \in K \subset V} J(v)$$

on écrit les conditions d'optimalité (conditions sur les dérivées de J).

Aussi très importants pour les algorithmes numériques (calcul du gradient de J).

Motivation: exemple en 1-d



Soit une fonction $f : [a, b] \to \mathbb{R}$ de classe C^1, avec $a < b$.

On veut résoudre

$$\inf_{x \in [a,b]} f(x)$$

Il est bien connu que les points de minimum de f se trouvent

- soit en un point x tel que $f'(x) = 0$,
- soit en a si $f'(a) \geq 0$,
- soit en b si $f'(b) \leq 0$.

(condition nécessaire mais pas suffisante !)

Différentiabilité

Soit V un espace vectoriel muni d'un produit scalaire $\langle u, v\rangle$ et d'une norme associée $\|u\|$.

Définition. On dit que la fonction J, définie sur un voisinage de $u \in V$ à valeurs dans $\mathbb{R}$, est différentiable au sens de Fréchet en u s'il existe une forme linéaire continue $L : V \to \mathbb{R}$ telle que

$$J(u+w) = J(u) + L(w) + o(w) \quad \text{avec} \quad \lim_{w \to 0} \frac{|o(w)|}{\|w\|} = 0 \, .$$

On appelle L la différentielle (ou la dérivée, ou le gradient) de J en u et on note $L = J'(u)$.

- ☞ Une forme linéaire L sur V est continue s'il existe $C > 0$ telle que $|L(w)| \leq C\|w\|$ pour tout $w \in V$.
- ☞ En dimension finie (par exemple, si $V = \mathbb{R}^N$) toute forme linéaire est continue, donc il n'y a rien à vérifier en ce qui concerne la continuité !

Remarques

☞ En dimension finie (et même dans un espace de Hilbert), on sait, grâce au théorème de représentation de Riesz , que pour toute forme linéaire continue L il existe un unique $p \in V$ tel que $\langle p, w \rangle = L(w)$. Par conséquent, on peut simplifier la définition en

$$J(u+w) = J(u) + \langle p, w \rangle + o(w) \quad \text{avec} \quad \lim_{w \to 0} \frac{|o(w)|}{\|w\|} = 0 \, .$$

☞ Pour calculer la dérivée (mais pas pour prouver la différentiabilité Fréchet !) on peut calculer la dérivée directionnelle. Pour $t \in \mathbb{R}$ et $u, v \in V$, on pose

$$j(t) = J(u + tv)$$

et on dérive $t \to j(t)$

$$j'(0) = J'(u)(v) = \langle J'(u), v \rangle.$$

Exercice

$x \in \mathbb{R}^n$ $\qquad J(x) = \frac{1}{2} Ax \cdot x - b \cdot x$ $\qquad b \in \mathbb{R}^n$

A matrice $n \times n$

A sym.

$$J(x+y) = J(x) + Ax \cdot y - b \cdot y + \underbrace{\frac{1}{2} Ay \cdot y}_{= o(y)}$$

$$= J(x) + (Ax - b) \cdot y + o(y) \Longrightarrow \boxed{J'(x) = Ax - b}$$

Changeons de produit scalaire : on remplace $\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$ par

$$\langle\langle x, y \rangle\rangle = \Pi x \cdot y \quad \text{avec } \Pi \text{ sym déf} \geq 0$$

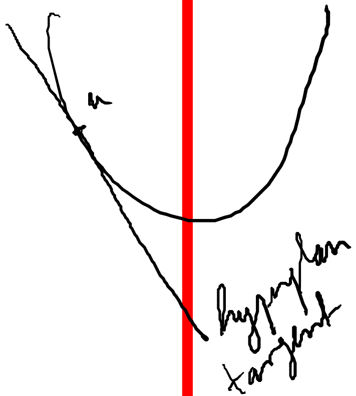
$$J(x+y) = J(x) + (Ax - b) \cdot y + o(y) \qquad = L(y)$$

$$(Ax-b)\cdot y = \Pi \Pi^{-1} (Ax - b) \cdot y = \langle\langle \Pi^{-1}(Ax - b), y \rangle\rangle$$

avec ce produit scalaire $\boxed{J'(x) = \Pi^{-1}(Ax - b)}$

Convexité

Proposition. Soit J une application différentiable de V dans $\mathbb{R}$. Les définitions suivantes de la convexité sont équivalentes



$$(1) \qquad J(\theta u + (1-\theta)v) \leq \theta J(u) + (1-\theta)J(v) \quad \forall\, u, v \in K \ ,\ \forall\, \theta \in [0,1]$$

$$(2) \qquad J(v) \geq J(u) + \langle J'(u), v-u \rangle \quad \forall\, u, v \in V$$

$$(3) \qquad \langle J'(u) - J'(v), u-v \rangle \geq 0 \quad \forall\, u, v \in V$$

monotonie

Preuve. $(1) \Rightarrow (2)$ $\quad J(\theta u + (1-\theta)v) = J(v + \theta(u-v)) = J(v) + \langle J'(v), \theta(u-v)\rangle + o(\theta(u-v))$

$\leq \theta J(u) + (1-\theta)J(v)$

$\Rightarrow \theta J(v) + \theta \langle J'(v), u-v\rangle + o(\theta(u-v)) \leq \theta J(u)$

on divise par θ, puis $\theta \longrightarrow 0$ $\Rightarrow$ $\boxed{J(v) + \langle J'(v), u-v \rangle \leq J(u)}$ (2)

$(2) \Rightarrow (3)$ $\quad J(u) + \langle J'(u), v-u \rangle \leq J(v)$

on somme $\Rightarrow$ $\langle J'(u) - J'(v), v-u \rangle \leq 0$

$t \in \mathbb{R}$
$s \in \mathbb{R}$

Suite de la preuve

Pour montrer (3) $\Rightarrow$ (1) on introduit $\phi(t) = J(u + t(v - u))$.

$\langle J'(u) - J'(v), u - v \rangle \geq 0 \quad (3)$

Dans (3) on remplace u par $u + t(v-u)$ et v par $u + s(v-u)$

$\langle J'(u+t(v-u)) - J'(u+s(v-u)), (t-s)(v-u) \rangle \geq 0 \quad (*)$

On $\phi'(t) = \langle J'(u+t(v-u)), v-u \rangle$

Donc $(*) \iff \left(\varphi'(t) - \varphi'(s) \right)(t-s) \geq 0$

Si $t \geq s$ alors $\varphi'(t) \geq \varphi'(s)$

on intègre en s entre 0 et θ $\Longrightarrow$ $\theta \varphi'(t) \geq \varphi(\theta) - \varphi(0)$

———— t entre 1 et θ $\Longrightarrow$ $\theta\left(\varphi(1) - \varphi(\theta)\right) \geq (1-\theta)\left(\varphi(\theta) - \varphi(0)\right)$

$\varphi(0) = J(u) \qquad \varphi(\theta) = J(\theta v + (1-\theta) u)$

$\varphi(1) = J(v)$

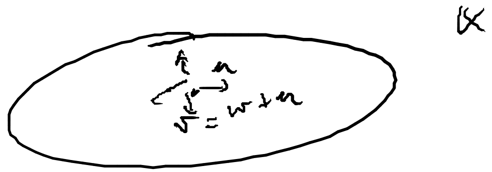
$\Longrightarrow \quad (1) \quad !$

Conditions d'optimalité: K convexe

Théorème (Inéquation d'Euler). Soit $u \in K$ convexe. On suppose que J est différentiable en u. Si u est un point de minimum local de J sur K, alors

$$\langle J'(u), v - u \rangle \geq 0 \quad \forall\, v \in K \ .$$

Réciproquement, si $u \in K$ vérifie cette inéquation et si J est convexe, alors u est un minimum global de J sur K.



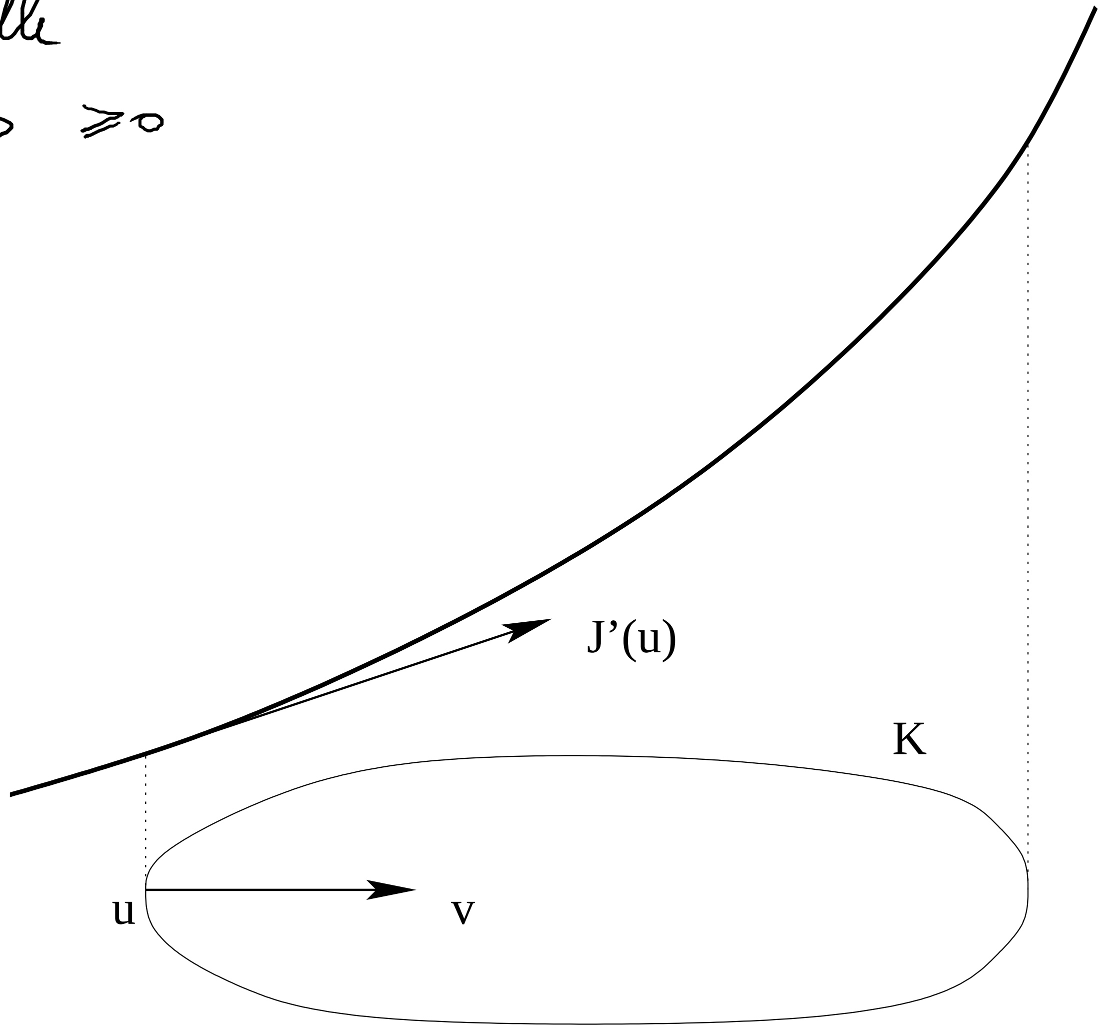
Remarques.

- ☞ Si u est intérieur à K, on obtient l'équation d'Euler $J'(u) = 0$.
- ☞ L'inéquation d'Euler est seulement nécessaire en général. Pour les fonctions convexes, elle est nécessaire et suffisante.

Inéquation d'Euler

dérivée directionnelle

$< J'(u), v-u > \geq 0$



La dérivée directionnelle de J dans la direction rentrante $(v - u)$ est positive.

Démonstration

Soit u un point de min

$$\cancel{J(u)} \le J(\underbrace{\theta v + (1-\theta)u}_{\in K}) \quad \forall v \in K \quad \forall\, 0 \le \theta \le 1 \quad \text{(convexité de } K\text{)}$$

$$\le J(u + \theta(v-u)) = \cancel{J(u)} + \langle J'(u), \theta(v-u)\rangle + o(\theta(v-u))$$

on divise par θ et ensuite $\theta \to 0$

$$\Longrightarrow \boxed{0 \le \langle J'(u), v-u\rangle}$$

Réciproque dans le cas où J est convexe :

$$J(\theta v + (1-\theta)u) \le \theta J(v) + (1-\theta) J(u)$$

$$\| $$

$$J(u) + \langle J'(u), \theta(v-u)\rangle + o(\theta(v-u)) \le \theta J(v) + (1-\theta) J(u)$$

$$\Rightarrow \theta J(u) + \theta \langle J'(u), v-u\rangle + o(\theta(v-u)) \le \theta J(v)$$

on divise par θ et on fait $\theta \to 0$

$$J(u) + \underbrace{\langle J'(u), v-u\rangle}_{\ge 0} \le J(v) \Longrightarrow J(u) \le J(v) \quad \forall v \in K$$

Exemple 1

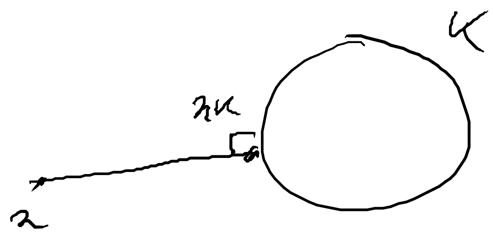
Projection sur un convexe fermé K. Pour $x \in V$, on cherche la projection orthogonale $x_K \in K$ de x sur K

$$\|x - x_K\| = \min_{y \in K} \|x - y\|.$$

La condition nécessaire et suffisante d'optimalité est

$$x_K \in K, \quad \langle x_K - x, y - x_K \rangle \geq 0 \quad \forall y \in K.$$

$J(y) = \|x - y\|^2$ $\qquad$ $J'(y) = 2\ (x - y)$



Exemple 1 bis

Projection sur le cône positif $K = (\mathbb{R}^+)^n$. Pour $x \in \mathbb{R}^n$, on cherche la projection orthogonale $x_+ \in K$ de x sur K

$$\|x - x_K\| = \min_{y \in K} \|x - y\|.$$

La condition nécessaire et suffisante d'optimalité est

$$x_K \in K, \quad \langle x_K - x, y - x_K \rangle \geq 0 \quad \forall y \in K.$$

On choisit y tel que $y_i = (x_K)_i$ pour tout $i \neq j$ et on en déduit

$$(x_K)_j = x_j \text{ si } x_j > 0, \qquad (x_K)_j = 0 \text{ si } x_j \leq 0$$

Exemple 2

Soit A une matrice carrée d'ordre n, symétrique définie positive. Soit B une matrice rectangulaire de taille $m \times n$ avec $m \leq n$. Soit b un vecteur de $\mathbb{R}^n$. On considère

$$\inf_{x \in \ker B} \left\{ J(x) = \frac{1}{2} Ax \cdot x - b \cdot x \right\}.$$

Il existe une unique solution $x^* \in \mathbb{R}^n$ qui vérifie

matrice transposée

$$Ax^* - b = B^* p \quad \text{avec } p \in \mathbb{R}^m.$$

avec p solution de $\underbrace{BA^{-1}B^*}_{m \times m} p = -BA^{-1}b$.

On appellera p le vecteur des multiplicateurs de Lagrange

Inéquation d'Euler (J étant convexe, c'est nécessaire et suffisant)

$\langle Ax^* - b, y - x^* \rangle \geq 0 \quad \forall y \in K = \ker B$

comme K est un e.v. $\quad y - x^* = z$ quelconque dans $\ker B$

$$< Ax^* - b, z > \geq 0 \quad \forall z \in \mathrm{Ker}\, B$$

on $-z$

$\Rightarrow \quad < Ax^* - b, z > = 0 \quad \forall z \in \mathrm{Ker}\, B$

$Ax^* - b \perp \mathrm{Ker}\, B$ Mais $(\mathrm{Ker}\, B)^\perp = \mathrm{Im}\, B^*$

$\Leftrightarrow \quad Ax^* - b \in \mathrm{Im}\, B^*$

Donc $\exists p \in \mathbb{R}^m$ tel que $Ax^* - b = B^* p$

Or $x^* \in \mathrm{Ker}\, B \Leftrightarrow Bx^* = 0 \Leftrightarrow B(A^{-1}b + A^{-1}B^* p) = 0$

$$\Leftrightarrow \boxed{BA^{-1}B^* p = -BA^{-1}b}$$

$BA^{-1}b \in \mathrm{Im}\, BA^{-1}B^*$

De plus si $\mathrm{rg}\, B = m \leq n \Rightarrow \mathrm{Ker}\,(BA^{-1}B^*) = \{0\}$

Exemple 3

Soit $L > 0$ et une fonction continue $f(x)$. On considère

$$\inf_{v \in V_0} \left\{ J(v) = \frac{1}{2} \int_0^L (v'(x))^2 \, dx - \int_0^L f(x) v(x) \, dx \right\}.$$

avec

$$V_0 = \left\{ v \in C^1[0, L] \text{ tel que } v(0) = v(L) = 0 \right\}.$$

La condition nécessaire et suffisante d'optimalité pour u est: trouver $u \in V_0$ tel que

$$\int_0^L u'(x) v'(x) \, dx = \int_0^L f(x) v(x) \, dx \quad \forall v \in V_0.$$

$$\begin{cases} -u'' = f & x \in (0, L) \\ u(0) = u(L) = 0 \end{cases}$$

Conditions d'optimalité: K non convexe

Exemples de K non convexe:

$$K = \{v \in V \text{ tel que } F(v) = 0\} \quad \text{ou} \quad K = \{v \in V \text{ tel que } F(v) \leq 0\}$$

Définition. Pour $v \in K$, on appelle cône des directions admissibles au point v

$$K(v) = \left\{ \begin{array}{l} w \in V \text{ tel que } \exists\, (v^n)_{n\geq 0} \in K\,,\ \exists\, (\varepsilon^n)_{n\geq 0} \in \mathbb{R}_+^* \text{ vérifiant} \\ \\ \lim_{n\to+\infty} v^n = v\,,\ \lim_{n\to+\infty} \varepsilon^n = 0\,,\ \lim_{n\to+\infty} \frac{v^n - v}{\varepsilon^n} = w \end{array} \right\}$$

Remarque. $0 \in K(v)$ et $K(v)$ est un cône car si $w \in K(v)$ et $\lambda \geq 0$, alors $\lambda w \in K(v)$.

Inéquation d'Euler (cas général)

Proposition. Soit u un minimum local de J sur K. Si J est différentiable en u, on a

$$\langle J'(u), w\rangle \geq 0 \quad \forall\, w \in K(u)\ .$$

Remarque. Tout le problème est d'identifier $K(u)$!

Preuve.

Exemples de cône $K(u)$

☞ Si u est intérieur à K, alors $K(u) = V$.

☞ Si K est convexe, alors $K(u) = \{w = v - u \text{ avec } v \in K\}$.

☞ Soit $F : V \to \mathbb{R}$ fonction régulière. Si $K = \{v \in V \text{ tel que } F(v) = 0\}$ et si $F'(u) \neq 0$, alors $K(u) = [F'(u)]^\perp$ (hyperplan tangent en u à la surface K).